ODF for Interoperability of Public Document in Indonesia

[1]I Made Wiryana, [2]Arrummaisha Adrifina, [3]Juwita Utami Putri, [4]Harya Iswara A.W

[1,2,3,4]Gunadarma University
[1,2,3,4]{mwiryana, arrum, jutami, harya}@staff.gunadarma.ac.id

Abstract

Open standards are widely considered to have a lot of benefits. This has led many governments to consider implementing open standards for document formats, such as the OpenDocument Format (ODF).

This paper will discuss aspects of the importance of proper document format for a document that must be long lasting and can be used by different organizations in the government. ODF is a document format that meet the criteria for office applications document storage. By moving to open standards, governments will see that they would benefit from choice, competition, and the ability to substitute different vendor implementations.

Keywords: government, Indonesia, interoperability, ODF, open standards

1 Introduction

Today, information and communication technology (ICT) architectures need to be flexible. They must be modular, plug-able and easy to set up, fast to integrate and fast to take down and re-purpose for governments and businesses alike to meet the demands of their citizens and customers. The architectures need to be built around agreed standard protocols and information needs to flow seamlessly across different applications and platforms.

A document format is a particular way that information is encoded for storage in a computer file. Document format can not be considered as simple or just follow what is usually used by people. Especially for documents belonging to government agencies that must be used by the public. Several technical aspects must be considered, such as:

- Open life-cycle
- Open availability
- Multiple implementation
- Interoperability across different systems
- Following a good document storage standard.

Besides, there are several considerations on non-technical aspects:

- Avoiding government agencies to "force" the public to buy certain word processing programs that can interact electronically. Because basically a document granted by a government agency to the public should be opened without any provision to use a certain type of software.
- Old documents can not be read because the vendor change its format in the new version of the program. Thus the selection of the document should also consider the age of the document for document preservation needs.
- Government agencies can not communicate because it uses a different version/product of word processor or using a format specified by the vendor. So it takes a standard document which can be used again by another government agency without being tied to one particular product type.

Public availability of documentation for early file formats varied. Some application vendors made their formats available on demand or published them in book form. Others released them under restrictive licenses that limited use by competitors. Through the years of 1980s and 1990s, office documents were primarily exchanged as printed copies. Where exchanges occurred electronically, they were most often between known parties, such as workers in the same office, using the same software. Electronic document exchanges was straightforward because the authors assumed that people whom they correspondence with are using the same version of software running in the same version of operating system. Interoperability among different software was difficult and trivial [5].

The difference between old style format and new style format is shown in figure 1. In old style format, information is closely linked to the application. Only the developers of these applications can understand the structure of the document. But the new trend has been started. One of the concepts

developed is the concept of document formats with open standards (open document format). Information is represented using a real open standard not under the control of a single vendor, and multiple applications can create and access it interchangeably. So that the document does not depend on an application, documents can be opened and edited using a variety of applications.

Old style New Style

Figure 1: Old and new style of documents[4]

Open Document Format (ODF) as one of the format of office application, supports information retrieval and exchange of documents without relating to the application or platform used when creating the document, either by using proprietary software and using Free/Open Source Software (FOSS). Many state governments around the world have adopted ODF, which is an open standard document format. What about the Indonesian government? Why the Indonesian government should start to adopt ODF? In this paper, we will look on document format selection, problem arised in selecting document format, what ODF is, why ODF, what benefits can be obtained, so that the Indonesian government can consider ODF as their standard document format for data exchange.

2 Approach in Selecting Document Format

Generally, document storage model is based on two spectrums, interface location (page image) and context/structure. Figure 2 shows us this condition

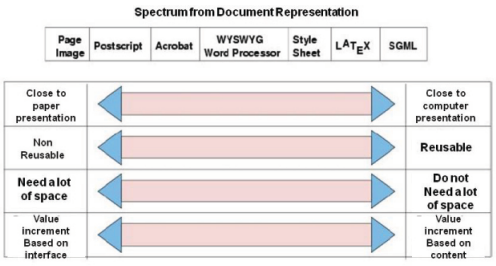

Figure 2: Spectrum of document storage

From figure 2 we can see that if someone want a reusable file, so the file model should be closed to SGML (Standardized General Mark-up Language). Figure 2 also explain why document in LaTex format is reusable if we compare it with wordprocesor in 1990's.

When the organization have to determine which file is used as a standard in all parts, they have to make a standardization of file types. Talk about a standard format, generally people do not refer to the same thing. There are two types of standard formats that are known in public[Mwiryana, 2010], they are:

- **De facto Standard.** This standard is something popular or used by many people in community. It is considered as a standard because it is known by many people, but not a standard from any institution and also untested. There is no guarantee from the application vendor to not to stop the support in future version since there is no standard from any legal institution. For example, program that support some format in their old version, do not support the old format in their new version. We can call this as a pragmatic approach in defining a standard for document format.

- **Defined Standard.** This standard specified in detail and associated with one standard organization (ex: ISO and ECMA). For open standard, every one can implement that standard free.

Government institution usually use de facto standard to determine document format because they thought every body have the application to open the document, but in reality is not true.

Beside that, from case study, document format that was choosen by de facto standard in year 1990's can not be opened with recent application. This example can be take as consideration in determine document that will be used in government institution.

Government institution, in using document to communicate with public, have to think about accessbility. Accessibilty is a chance for people to access the document in public area. Because of that, public may not be bound with some program to open and edit the document.

3 Problem of Pragmatic Approach

Problem of pragmatic approach in which technical appraisal of the material is used to preserving one format over another. These approach provides for nine principles to consider as part of appraisal.

These principles balance economically sustainable preservation and intellectual 'value' with the practicalities of working with specific, and especially proprietary, file formats.

Technical factors are only part of the reality assessment of material for inclusion in collections. Other factors such as intellectual content, significance of the document, significance of the donor/creator and any relationship to material already also play a part. Document considers 'original' formats accepted for long-term preservation, and does not consider formats appropriate for dissemination.

A Company or Nation looking to preserve born digital document permanently and have little control over the formats in which document is transferred or deposited. Nowadays internet membership succeed to connect people all over the world, there is a change in the document distribution pattern. Authors can now easily upload documents in any format to the next can be downloaded by anyone. Each documents format can be opened only by using certain software. When the person downloading the file, author could not ensure that the receiving party was using the same software3.

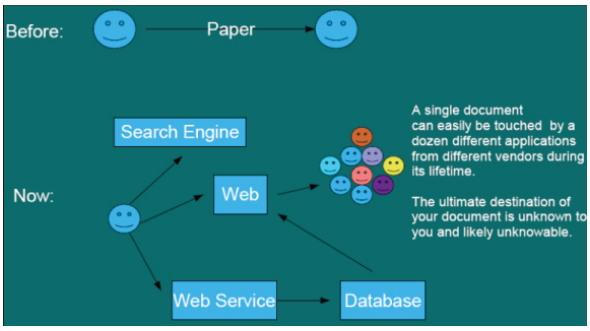

Figure 3: Before and After Internet Development

The ideal intervention point from a preservation perspective is at the point digital material is first created. However this may be unrealistic. Many working within organizations have no choice in the applications they use, cost of applications may be an issue, or there may simply be a limited number of applications available on which to perform specialist tasks. It may be obsolete, in obscure formats, on obsolete media and without any metadata describing its context, creation or rendering environment.

Computer applications keep their data in formats, each application typically having its own file format. The Web site filext lists some 25,000 file extensions in its database.

The long-term preservation of any format depends on the type of format, issues of obsolescence, and availability of hardware and/or software, resources, experience and expertise. This presents a number of problems:

1. the proliferation of file formats;
2. the use of proprietary file formats, and
3. formats becoming obsolete, either by being incompatible with later versions of the applications that created them, or by those applications no longer existing.

This assumes that proprietary formats are more problematic to preserve as their structure and composition are not known, which hinders preservation intervention by imposing the necessity for specialist expertise. Moreover, as new software is created, so new file formats proliferate, and consequently exacerbate the problem. The primary way to solve this is by creating and adopting standards. This is what makes the issue of interoperability become important. A good document format will facilitate the achievement of interoperability in the use of software. With the method and good storage format, then a file will be used by various software.

Acceptability of any format is based on each person's ability to manage and preserve that format successfully into the future, set against the intellectual 'worth' of that material. Florida Digital Archive uses nine principles to evaluate the acceptability, or otherwise, of specific formats.

Principle
Formats in current, widespread or common use
Formats which are non-proprietary
Formats which are standards-based
Formats for which specifications are publicly available
Formats which offer platform-independence
Formats which are uncompressed
Considerations of the intellectual content of the file(s) and/or the importance of the creator
Considerations of how material complements existing material
Economic/resource implications

Cost and the resources required to work with obsolete media and files will play a role in the appraisal process, although cost alone will not be a factor in the decision to accept document. Subjective judgement will be applied to the provenance of the document and the likelihood that not only can document be recovered but that it will be historically significant and worthy of the effort of preservation. Example, the impact of user locked-in by

application vendor and the deleterious effects that ensue. If a document format is designed closely reflect the internal data structures of a single vendor's applications, then documents created using those applications will only work, or work well, with those applications. Some vendors often make the development and update of applications that already exist. This will impact on old documents which sometimes can not be read because the vendor to change its format in the new version of the program. Thus the selection of the document should also consider the age of the document for document preservation needs. As the user continues to create and accumulate documents in that proprietary format, his or her ability to evaluate and migrate to alternative PPAs in the market diminishes due to substantial switching costs. The presence of high switching costs is the essence of vendor lock-in and is a significant challenge from the perspective of the user, who now faces effectively diminished options; the potential competitor, who finds it more difficult to enter a market dominated by an incumbent with an established lock-in; and public policy, which has a long-standing practice of encouraging competition. One fundamental way to reduce the strength of vendor lock-in and encourage substitutability of equivalent goods is by adopting and using open standards that encourage interoperability.

4 Open Document Format (ODF)

The Open Document Format for Office Applications (also known as OpenDocument or ODF) is an XML-based file format for representing electronic documents such as spreadsheets, charts, presentations and word processing documents. The specifications were originally developed by Sun Microsystems and the standard was developed by the OASIS ODF TC, committee of the Organization for the Advancement of Structured Information Standards (OASIS) consortium and based on the XML format originally created and implemented by the OpenOffice.org office suite [6].

The most common filename extensions used for OpenDocument documents are:

- .odt for word processing (text) documents
- .ods for spreadsheets
- .odp for presentations
- .odb for databases
- .odg for graphics
- .odf for formula, mathematical equations.

At the beginning, ODF was only used for office application, OpenOffice. Then the Technical Committee of OASIS developed this document format. In May 2005 the OASIS consortium approved this format to become a standard and in May 2006 approved the Open Document format as an international standard, ISO / IEC 26300:2006. Figure 4 shows us the ODF development from 2005 to 2010.

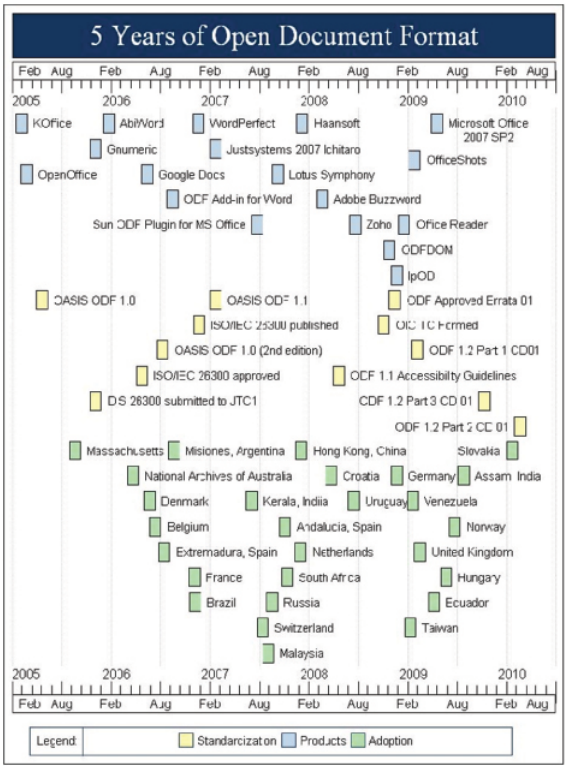

Figure 4: ODF development

Further developments of ODF version are:

- OpenDocument 1.0 (second edition) includes the editorial changes made. It is available in ODF, HTML and PDF formats.
- OpenDocument 1.1 includes additional features to address accessibility concerns. [10]
- OpenDocument 1.2 includes additional accessibility features, RDF-based metadata, a spreadsheet formula specification based on OpenFormula, support for digital signatures and some features suggested by the public.

Many countries has adopted OpenDocument as their national standards[1], such as:

- **North America**
  (especially United State of America). For example in the state of Massachusetts which, although not yet requiring only the ODF, but already include consideration of the document standard for office applications

- **Latin America**
  Brazil (ePING 2006), Venezuela, Argentina and Uruguay. Countries such as Brazil and even issued a rule equivalent to the Act for the use of ODF in government agencies.

- **Europe**
  In France, through Decree 2007-284, March 2007 ODF must be used in government agencies. Through the Belgian Council of Ministers decision of 2006, Belgia also requires the use of ODF. Croatia, Norway (via eNorge 2009) requires that the use of ODF document standard one of which is mandatory. Italy make ODF as a national standard and must be used within government agencies. In Spain (in the state of Andalucia and Extremadura), also Netherlands, ODF is the standard for reading, publishing and the exchange of information for all governmental organisations and the government agencies are required to use ODF.

- **Asia**
  Malaysia, and India (some states and companies).

5 Interoperability of Service

Interoperability can be reached by some aspects:

- **Open Document Format**, means someone can open a file which was created by someone else from different application.

- **Open Exchange Format**, in this level, any of application system can exchange their data with other applications. For example the Human Resource Information System in institution A can do a data exchange with the Human Resource System in institution B.

- **Open Service**, in this level, many applications can interact each other by providing an open service. Other system can request for a service with open mechanism without being tied by certain platform, data format, or hardware.

Figure 5 shows the interoperability of open data.

Sam Hiser in [3] analyzed the "openness" of two emerging XML-based document formats, which are ODF and Office Open XML (OOXML) from Microsoft Corporation. ODF was revealed as open across four criterias (open life-cycle, open availability, multiple implementation, and interoperability across different systems) when compared to OOXML.

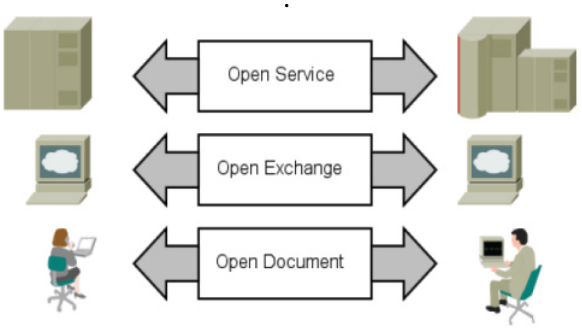

Figure 5: Interoperability

6 The Importance of ODF for Indonesian Government

Some explanations above clearly describe some importance of ODF, especially for government, such as:

1. Document format standardization in government agencies in Indonesia, to make easier the use of information among government agencies. Therefore, public documents can be opened and used by any different application and platform.

2. The interoperability of documents is needed not only among government agencies, but also among government agencies and citizens. Anyone (citizen) who needs data and information from the government can open the documents with many applications. ODF can be opened from many applications, such as: OpenOffice.org, KWord, NeoOffice, StarOffice, and many others [7]. This also reduce the dependency with certain vendor.

3. Avoid to lose information in the future. Government agencies need to preserve their data so they can access, modify, and save archived documents in the future. In 1990s, Wordstar and Word Perfect were famous as an office application, as well as Lotus 1-2-3 for spreadsheet and DB-Three for database application. At that time, governments used them to process their documents. Now, it is difficult if they would like to read those documents. ODF, as an open, durable standard, can ensure that a government document saved today will not be locked by any platform or abandoned tomorrow. A government should never be dependent on a single vendor's technology to use its own information.

4. Low cost implementation. To access the open document format (document with open standard) doesn't need a royalty fees. Government can use free/open source software, so

the state funds can be used for other needs in the areas of government, industry, education without having to pay licensing fees. Imagine how much the cost would be to buy Microsoft Office 2007 if the government used the docx format. Because previous versions of MS Office can not be used to access the data.

Considering those aspects, Indonesia is moving to implement ODF as a national standard. ODF is now being associated with the Indonesian National Standard (SNI). It is hoped that this year the government can finish the standardization and next year they can socialize to migrate all government's important document to ODF. The ministry of communication and informatics also advices private sectors who serve public to do an ODF migration in order to give access to public easily without having any dependence to certain vendor. [2]

7 Conclusion

Indonesian government has a problem in document interoperabilty among their departments. Usually, each department have their own standard of document format. This conditions cause a problem in communication and document exchange between department.

Open Document Format or ODF, is a solution of document interoperability among the departments. Because it has some advantages, they are : ODF can be opened by all kinds of application, lower cost, long life cycle and information ownership is cleared. Because of these advantages, many countries already implement ODF as their standard of document format.So, now is the time for indonesian government to implement ODF as standard of our document format

References

[1] Odf alliance. 2008. http://www.odfalliance.org/resources/Adoptions-ODF-Dec2008.pdf.

[2] Detikinet dokumen negara wajib gunakan format odf. 2010. http://www.detikinet.com/read/2010/07/28/151112/1408538/398/dokumen-negara-wajib-gunakan-format-odf/.

[3] S. Hiser. Achieving openness: a closer look at odf & ooxml. odf alliance. 2007.

[4] M. Marcich. Open document format. it's your information. 2008. http://www.odfalliance.org/resources/Adoptions-ODF-Dec2008.pdf.

[5] R. Weir. Opendocument format the standard for office documents. *IEEE Computer Society*, 2009.

[6] Wikipedia. Opendocument. 2010. http://en.wikipedia.org/wiki/OpenDocument.

[7] Wikipedia. Opendocument software. 2010. http://en.wikipedia.org/wiki/OpenDocument_software.